University of Groningen

A Condorcet Jury Theorem for Unknown Juror Competence

Atkinson, D.; Romeijn, J.-W.

Published in:
Mathematics

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2009

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Atkinson, D., & Romeijn, J-W. (2009). A Condorcet Jury Theorem for Unknown Juror Competence. *Mathematics*.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 05-05-2023

A Condorcet Jury Theorem for Unknown Juror Competence

Jan-Willem Romeijn and David Atkinson
Department of Philosophy
University of Groningen
j.w.romeijn@rug.nl d.atkinson@rug.nl

Abstract

This paper presents a generalisation of the Condorcet Jury theorem by relaxing the assumption that the competence of the jurors is fixed. Instead we assume a uniform prior probability assignment over the possible competences, adapt this assignment in the light of the jury vote, and then compute the posterior probability, conditional on the jury vote, of the hypothesis voted over. One of the more notable aspects of this posterior probability is that it depends on the size of the jury as well as on the absolute margin of the majority.

1 Condorcet revisited

In this section we introduce Condorcet's jury theorem. We also present the results of List [2004], according to which the posterior probability of the hypothesis voted over, conditional on the jury vote, only depends on the absolute margin of votes in favour and against the hypothesis.

Let H^1 be the hypothesis that Jack murdered Jill and H^0 the hypothesis that he did not, so $\{H^0, H^1\}$ is a partition. Suppose that of a jury of n members trying Jack, a number n_1 vote that H^1 is true, while the remaining n_0 members vote that H^0 is true. For both $j = 0$ and $j = 1$, we assume that if H^j is in fact true, the probability that jury member i votes for H^j being true, denoted by V_i^j, is some fixed chance q_j, which we will call the competence of the jurors on the hypothesis H^j. For all $i, i' = 1, \ldots, n$ and

$j, j' = 0, 1$, if $i' \neq i$, we set

$$p(V_i^j | H^j \cap V_{i'}^{j'}) = p(V_i^j | H^j) = q_j ,$$

The left equality says that the jurors all vote independently, and the right one that they vote with fixed competences q_j, for both $j = 0, 1$. Note that the competences of the jurors with respect to H^0 and H^1 do not refer to a general ability to judge. They are specific for the hypotheses, and the competences for the hypotheses H^0 and H^1 can differ: jurors might be more accurate in judging the former than in judging the latter. Finally, we assume that the competences will be greater than one half, $q_j > 1/2$, so the judgment of jury members is better than the result of tossing a fair coin.

We can now introduce Condorcet's jury theorem. Say that Jack is indeed guilty, H^1, so that the probability for any jury member to vote in favour of Jack's guilt, V_i^1, namely q_1, is greater than one-half. Now for an ever larger jury size n, consider the relative frequency of voters in favour, $f_1 = \frac{n_1}{n} = 1 - f_0$. By the law of large numbers, the probability that f_1 differs from q_1 tends to 0. Because $q_1 > 1/2$, we have that the probability of a correct majority vote, $n_1 > n_0$, tends to 1 in the limit. We refer to Dietrich [unpublished] for the proof of a somewhat more general version of this theorem.

List [2004] emphasises the importance of the converse of this result. Rather than calculating the probability of a majority of votes V_i^j given the truth of H^j, we want to know the probability of correctness of the hypothesis H^j, given a majority of votes V_i^j. For convenience we denote the difference between votes for H^1 and H^0, also called the absolute margin of the majority, by $\Delta = n_1 - n_0$. We denote the entire jury vote by $V_{n\Delta} = \cap_{i=1}^n V_i^{u(i)}$. Here $u(i) = 1$ if jury member i voted for H^1 and $u(i) = 0$ if she voted for H^0, so that $n_1 = \sum_i u(i)$ and $n_0 = n - n_1$. By Bayes' theorem we have

$$\begin{aligned} p(H^1 | V_{n\Delta}) &= \frac{p(\cap_i V_i^{u(i)} | H^1) p(H^1)}{p(V_{n\Delta})} \\ &= \frac{\prod_i p(V_i^{u(i)} | H^1) p(H^1)}{p(V_{n\Delta})} \\ &= \frac{q_1^{n_1} (1 - q_1)^{n_0} p(H^1)}{p(V_{n\Delta})}. \end{aligned}$$

where $p(H^1)$ is the prior probability of the hypothesis H^1. For H^0 we can derive a similar expression, replacing q_1 by q_0 and swapping the roles of

n_0 and n_1. The denominator $p(V_{n\Delta}) = p(\cap_i V_i^{u(i)})$ is the same in both equations.

We can avoid calculating the denominator $p(V_{n\Delta})$ by using the posterior odds instead of the posterior probability:

$$\frac{p(H^1|V_{n\Delta})}{p(H^0|V_{n\Delta})} = \frac{q_1^{n_1}(1-q_1)^{n_0}p(H^1)}{q_0^{n_0}(1-q_0)^{n_1}p(H^0)}. \tag{1}$$

This is the odds that Jack is guilty, given the jury verdict. Since $n_0 = {}^{(n-\Delta)}/_2$ and $n_1 = {}^{(n+\Delta)}/_2$, the posterior odds depends both on the absolute margin Δ and on the jury size n.

For the posterior odds we can derive an inverse variant of Condorcet's theorem. If we let the jury size n go to infinity, and assume a fixed relative frequency $f_1 > {}^1/_2$, the posterior odds for H^1 will tend to infinity as well. This is a direct consequence of the fact that the function

$$h_n(q) = q^{nf_1}(1-q)^{n(1-f_1)}$$

has a peak at $q = f_1$, and the further fact that this peak gets sharper for larger n. Hence the likelihood ratio

$$\frac{h_n(q_1)}{h_n(1-q_0)} = \frac{q_1^{nf_1}(1-q_1)^{n(1-f_1)}}{q_0^{n(1-f_1)}(1-q_0)^{nf_1}}$$

will tend to infinity in the limit of n going to infinity. This is independent of the values that we fill in for the competences $q_j > {}^1/_2$.

By making two further assumptions we can arrive at the central result in List [2004], the so-called Condorcet formula. First, we assume that the priors of H^0 and H^1 are equal, $p(H^1) = p(H^0)$, although this is not a crucial assumption. It means that initially we judge it equally probable that Jack committed the murder as that he did not. Second, and this is crucial for List's result, we assume that the competences of jury members on H^0 and H^1 are equal, $q_0 = q_1 = q$, meaning that all jury members are precisely as reliable in condemning a murderer as they are in acquitting an innocent suspect. With these two assumptions Equation (1) simplifies to the Condorcet formula

$$\frac{p(H^1|V_{n\Delta})}{p(H^0|V_{n\Delta})} = \left(\frac{q}{1-q}\right)^{\Delta}. \tag{2}$$

On the assumption that $q > {}^1/_2$, we have that $\frac{q}{1-q} > 1$ so that the posterior odds that Jack killed Jill is larger than 1 if $\Delta > 0$ and smaller than 1 if

$\Delta < 0$. The posterior odds depends only on the absolute margin between the numbers of correct and incorrect votes brought out by the jury members, and not on their total number. Note that this is perfectly consistent with both the Condorcet theorem and its inverse version for posterior odds: for increasing jury size n and fixed competence q_1, the expected value of Δ increases with n, and for increasing jury size n and fixed relative frequency f_1, the fixed value of Δ increases with n.

Now focus on the fact that, for a given Δ, the posterior odds do not depend on the jury size. Imagine there are two juries, one with 10 members and one with 100 members. Suppose that both juries vote on the guilt of Jack, and that the 10-member jury unanimously votes for guilt while the 100-member jury votes by 56 in favour, and 44 against Jack's guilt. Which of the two juries then makes the guilt of Jack more probable? Well, the majority in the former is less than the majority in the latter, i.e. $\Delta_{10} = 10 < 12 = \Delta_{100}$, so that, according to Equation (2), the probability of Jack's guilt is greater for the larger than for the smaller of the two juries. Hence, if we want to have as much certainty as we can get, apparently we should prefer the verdict of the larger jury.

As we will argue below, there is something suspect in this conclusion. Indeed, we surmise that Equation (2) is too strong an idealisation. It is based on the unwarranted assumption of a symmetrical and fixed competence. By considering a model in which both these assumptions are dropped we show the following:

- The probability that the jury majority verdict is incorrect is monotonically increasing in the jury size n, if the absolute margin Δ remains constant.
- The probability that the jury majority verdict is incorrect tends to one-half as n tends to infinity, if Δ remains constant in this limit.
- The probability that the jury majority verdict is incorrect tends to zero as n tends to infinity, if the fractional majority, $f = \Delta/n$, tends to a nonzero constant in this limit.

The exclusive dependence on the absolute margin is thus seen to be an artefact of idealising assumptions, and not something inherent to real jury verdicts. The inverse Condorcet jury theorem, on the other hand, will prove to be valid in the new model as well.

2 A counterintuitive consequence

In this section we argue, by means of a classical statistical analysis, that there is something rather puzzling about the Condorcet formula (2), according to which only the absolute margin matters when one assesses the probability that a jury vote lends to the hypothesis voted over. The problem is that no account has been taken of how probable the jury votes are to begin with.

We first make precise the problem with the Condorcet formula by constructing a confidence interval for juror competence that depends on majority and jury size. As in the previous section we assume that H^1 is true and that the competence parameter is q. Each juror votes independently and with identical probability, so that the number of votes n_1 has a binomial probability distribution. Its expectation is $\mathrm{E}[n_1] = nq$, and the standard deviation is $\mathrm{SD}[n_1] = \sqrt{nq(1-q)}$. So the mean and standard deviation of the majority Δ are

$$\begin{aligned} \mathrm{E}[\Delta] &= \mathrm{E}[n_1] - \mathrm{E}[n_0] = 2\mathrm{E}[n_1] - n = n(2q-1)\,, \\ \mathrm{SD}[\Delta] &= 2\sqrt{nq(1-q)}\,. \end{aligned}$$

For any given competence q and jury size n, we have a probability of roughly 95% that Δ lies within the specific bounds of two standard deviations around the mean, $\mathrm{E}[\Delta] - 2\mathrm{SD}[\Delta] < \Delta < \mathrm{E}[\Delta] + 2\mathrm{SD}[\Delta]$.

On this basis we can construct a confidence interval for q. Suppose Δ lies at the edge of the interval indicated above. Then we would have one or other of the following:

$$\Delta = n(2q-1) \pm 4\sqrt{nq(1-q)}\,,$$

which can be solved for q as a function of Δ, yielding the two roots

$$q_{\min}\,,\, q_{\max} = \frac{1}{2(n+4)}\left(n + 4 + \Delta \pm \sqrt{n + 4 - \frac{\Delta^2}{n}}\right)\,.$$

By way of interpretation, we say that any competence q from within the interval $[q_{\min}\,,\, q_{\max}]$ entails that the given majority Δ and jury size n are not so improbable that they are a cause for worry. If the $q_{\min}$ of a jury characterised by n and Δ is greater than the $q_{\max}$ of a jury with n' and Δ',

$$q_{\min}(n, \Delta) > q_{\max}(n', \Delta')\,,$$

then we can say that something very improbable has occurred: at least one of the two jury votes has in that case voted out of the ordinary. This would give us cause to reconsider the assumptions of the statistical model at issue.

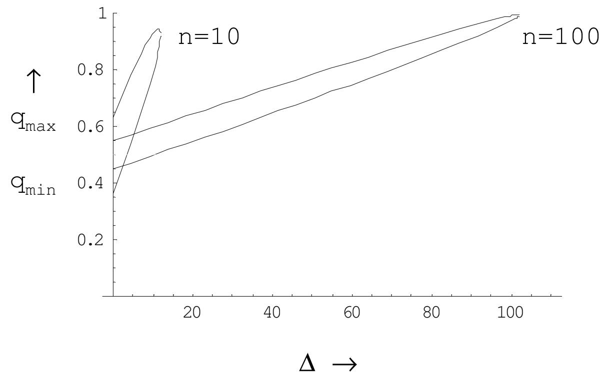

Figure 1: A graph of the bandwidth of reasonable values of the competence q against the jury majority Δ, for both $n = 10$ and $n = 100$.

Now let us have a second look at the example provided in Section 1. In Figure 2 the extremal values of q are plotted against the majority Δ for the two juries of 10 and of 100 members. For $\Delta_{10} = 10$ it turns out that $q_{\min}(10, 10) = 0.786$, while for $\Delta_{100} = 12$ we find $q_{\max}(100, 12) = 0.606$. Moreover, since $q_{\max}(100, 50) = 0.783$ and $q_{\max}(100, 52) = 0.792$, we need a majority of at least 76 against 24 in the jury of 100 to feel that there is no cause for worry. With the given votes, we must therefore conclude that something highly improbable did occur.

Now perhaps we simply know the numerical value of the juror competence. Or perhaps we know that all jurors have equal competence without knowing its value, in which case we might say that the competence of the jury is larger than what is suggested by the larger jury, or that the competence of the jurors is much smaller than what is suggested by the smaller jury, or possibly both. In any such case, the result of List is applicable, and we must simply conclude that we have witnessed a freak accident.

Alternatively, we might conjecture that the two juries have different values for the juror competence. In order to sort this out, we can formulate the hypothesis that the jurors from the smaller jury are in fact more competent, and perform a statistical test on this. But the idea that the jury vote tells us something about the competence of the jurors can also be taken a step further. We can say that the vote of the jury indicates the competence of the jury directly, and that this may be used with advantage in the choice between jury verdicts. Specifically, the unanimous vote of the jury of 10 should perhaps weigh more heavily, despite the rule of Equation (2), simply because the unanimity suggests that the jurors are competent.

In other words, the suggestion here is that a jury vote reflects more than just the truth or falsity of the hypothesis voted over. It also conveys information on how easy it is for jurors to vote correctly. A close call in the jury, such as the small majority of 12 in the jury of 100 members, indicates that the jurors find it hard to tell whether Jack murdered Jill, while the unanimous vote of the small jury seems to suggest that the jurors find Jack's guilt fairly clear. We will devote the remainder of this paper to making precise what the size of the majority tells us about the competence of the jurors, and what the consequences of that are for assessing the jury vote.

Before starting on this, however, we briefly consider some alternative explanations for the jury votes considered above. First of all, List [2004] shows that if we do not make the assumption of symmetric competence, $q_0 = q_1$, but instead let these competences vary independently, the posterior odds do depend on the jury size. Depending on which of the two competences is larger, a larger jury with equal absolute margin will have smaller or larger posterior odds for the hypothesis. By choosing the competences in the right way, this effect may even cause the posterior odds to lean towards the minority vote. However, the model with differing but fixed competences q_0 and q_1 fails to capture the intuitions on jury verdicts voiced above. It introduces a dependence on jury size of an entirely different nature, one that is not related to our present concerns.

Another possible extension of the model of jury decisions fares better in this respect. Arguably, a small jury has a completely different group dynamics than a larger jury, and the jury verdict may reflect how the jurors have interacted. For example, it may be that jurors adjust their views to coincide with those of jurors sitting in close proximity to them. In a jury of 100 the jurors may then still be treated as approximately independent. But in a jury of 10, all jurors are in close proximity to each other. Therefore a unanimous vote in a small jury may very well be the result of mindless groupthink rather than of high juror competence. Such failures of independence will generally throw the results of a jury vote into a different perspective.

In a similar vain, we may think that the coherence of the jurors in the smaller jury is indicative of the veracity of the jury verdict. This idea is at the basis of the discussion that Bovens and Hartmann [2004] give of the Condorcet formula. They note its counterintuitive consequences, adapt the model to include a positive correlation between the votes, and then show that in this model a smaller unanimous jury lends more credibility to the jury verdict than a larger jury whose verdict is divided, even while the absolute margins in both juries are equal. Moreover, by adapting the parameters in the model they can vary the degree to which the coherence of jurors adds to the credibility of the verdict.

The coherence model of Bovens and Hartmann provides a successful explication of certain intuitions concerning jury votes. It sensibly drops the assumption of the independence of the jurors, and employs the truth-conduciveness of the coherence of votes to avoid the counterintuitive consequence of the Condorcet formula. A drawback of this solution is that it relies on particular parameter values that must be filled in at the start. Given these parameters, we can deduce the dependence of the posterior odds on the jury size, but this dependence is in a sense put in by hand. But this drawback does not mean that we should discard the coherence model.

Accordingly, we do not motivate the model of the following sections with the fact that it captures our intuitions on jury votes better. Rather it captures other intuitions about jury votes, differing from those captured in the coherence model, namely that the competence of jurors can be revealed by the jury vote. We think that these intuitions are of interest in their own right. In addition, we think that the resulting model meshes better with the structure revealed by the classical statistical analysis of the preceding section.

3 Jury vote with unknown competence

In the following we present a model in which the jury vote is indicative of how competent the jurors are concerning the hypothesis at hand. We retain the assumption that the jurors vote independently and concentrate on relaxing the assumption of a fixed juror competence. To do so we employ Bayesian statistical inference. We first compute a posterior probability assignment over the competences q_0 and q_1 for H^0 and H^1 respectively, based on the given jury vote and a prior probability over competences and hypotheses. This inference determines how the jury vote informs us of the competence: we may derive an expectation value for the juror competences from it. More importantly, we compute the probability that a jury vote gives to the hypothesis voted over.

Computing the expectation value for the jury competence is a tricky business. In the foregoing we had a partition of two hypotheses, H^0 and H^1. But since the competence parameter is unknown, we must split these hypotheses up into ranges of hypotheses, H_{q_0} and H_{q_1}. The expressions $p(H_{q_j})$ should therefore be regarded as probability densities rather than themselves probabilities. The hypotheses H^0 and H^1 each consist of a range of statistical hypotheses, parameterised by q_0 and q_1 respectively. These hypotheses have the likelihoods

$$p(V^i_j | H_{q_j} \cap V^{i'}_{j'}) = q_j$$

for $j = 0, 1$. The probabilities of the aggregate hypotheses H^0 and H^1 are

$$\begin{aligned} p(H^0) &= \int_0^1 p(H_{q_0})\, dq_0 \,, \\ p(H^1) &= \int_0^1 p(H_{q_1})\, dq_1 \,. \end{aligned}$$

Further, we assume that the prior is equal and uniform over the interval $(\frac{1}{2}, 1)$, for both q_0 and q_1, meaning that $p(H_{q_j}) = 1$ for $\frac{1}{2} \leq q_j \leq 1$, and $p(H_{q_j}) = 0$ for $0 \leq q_j < \frac{1}{2}$. Thus the only prior assumption is that the jury members are not incompetent, but aside from that the prior density is flat. The above considerations entail

$$p(H^0) = \int_{1/2}^1 1\, dq_0 = \frac{1}{2} = \int_{1/2}^1 1\, dq_1 = p(H^1).$$

For reasons of simplicity we will not deviate from this assumption in what follows.

It is convenient to reduce the number of parameters in this statistical model to a single one by a suitable substitution of the parameters over the domain. In the above setup, the use of the two parameters q_0 and q_1 does not mean that the statistical model is two-dimensional. The likelihoods involve q_0 if H^0 is true and q_1 if H^1 is true, but these are mutually exclusive hypotheses, so there is no overlap in which the likelihoods involve both parameters. Because of this we can employ a single range of hypotheses H_r with the parameter domain $r \in [0,1]$, which is formally equivalent to the combination of q_0 in H^0 and q_1 in H^1.

Let us make this formal equivalence precise. First, within the domain $r \in [0, 1/2)$, and setting $q_0 = 1 - r$, we have the following equalities:

$$\begin{array}{rcccccl} p(V_i^0|H_{q_0}) & = & q_0 & = & 1-r & = & p(V_i^0|H_r) \\ p(V_i^1|H_{q_0}) & = & 1-q_0 & = & r & = & p(V_i^1|H_r)\,. \end{array} \tag{3}$$

Similarly, in the domain $r \in (1/2, 1]$, and setting $q_1 = r$, we have the following equalities:

$$\begin{array}{rcccccl} p(V_i^0|H_{q_1}) & = & 1-q_1 & = & 1-r & = & p(V_i^0|H_r) \\ p(V_i^1|H_{q_1}) & = & q_1 & = & r & = & p(V_i^1|H_r)\,. \end{array} \tag{4}$$

In words, there is a formal equivalence between the likelihoods of the hypotheses H_r for $r < 1/2$, and those of H_{q_0} for $q_0 > 1/2$. Similarly, there is an equivalence between the likelihoods of the hypotheses H_r for $r > 1/2$ and those of H_{q_1} for $q_1 > 1/2$.

Now consider the resulting likelihoods for the hypotheses H_r. From the right hand side of Equations (3) and (4) we can see that, over the entire domain $r \in [0,1]$, the hypotheses H_r have the likelihoods

$$\begin{array}{rcl} p(V_i^0|H_r) & = & 1-r\,, \\ p(V_i^1|H_r) & = & r\,. \end{array}$$

By updating the separate hypotheses H_r according to these likelihoods, we are effectively updating the hypotheses H_{q_0} and H_{q_1} for each of the values $q_0 \in (1/2, 1]$ and $q_1 \in (1/2, 1]$.

Next we consider the priors over the hypotheses H_{q_0} and H_{q_1}. Recall that we assumed a uniform prior probability distribution over both of them. But we can rewrite these priors in terms of priors over the hypotheses H_r in the respective domains $r \in [0, 1/2)$ and $r \in (1/2, 1]$, as follows:

$$\begin{array}{rcl} p(H^0) & = & \displaystyle\int_{1/2}^{1} p(H_{q_0})\, dq_0 = \int_{0}^{1/2} p(H_r)\, dr\,, \\ p(H^1) & = & \displaystyle\int_{1/2}^{1} p(H_{q_1})\, dq_1 = \int_{1/2}^{1} p(H_r)\, dr\,. \end{array}$$

Hence the uniform priors over the hypotheses H_{q_0} and H_{q_1} translate into a single uniform prior over the hypotheses H_r with $r \in [0,1]$. We can interpret the probability of H_r with $r < 1/2$ as the probability of H_{q_0} by the translation $q_0 = 1 - r$, and similarly, we can interpret the probability of H_r with $r > 1/2$ as the probability of H_{q_1} by the translation $q_1 = r$. We note, as an aside, that it is attractive to start out with uniform priors over the hypotheses H_{q_j}, or at least with priors that combine into a Beta distribution over H_r. Priors over H_{q_j} that have a different shape do not necessarily lead to posterior distributions that can be expressed analytically.

The substitution above is useful because we have thereby disposed of a parameter, replacing q_0 and q_1 by the single parameter r. Moreover, we can model the impact of the jury vote on the combined uniform probability assignments over $q_0 \in (1/2, 1]$ and $q_1 \in [1/2, 1]$ by modelling its impact on the uniform probability assignment over $r \in [0, 1]$.

We shall condition this distribution on the jury vote $V_{n\Delta}$, characterised by the numbers of votes n_0 for H^0 and n_1 for H^1, or equivalently, by the size of the jury $n = n_1 + n_0$ and the majority $\Delta = n_1 - n_0$. Then the posterior probability distribution over H_r results in a well-known form for the posterior distribution, the Beta distribution,

$$p(H_r|V_{n\Delta}) = \frac{(n+1)!}{n_0!n_1!} r^{n_1}(1-r)^{n_0},$$

with $r \in [0, 1]$. For $r > 1/2$ we are thereby indirectly specifying the posterior probability distribution over the hypotheses H_{q_1} according to the transformation $q_1 = r$, while for $r < 1/2$ we are indirectly specifying the posterior for the hypotheses H_{q_0}, using the transformation $q_0 = 1 - r$.

From this expression we can derive the posterior probability of the hypotheses H^0 and H^1:

$$p(H^0|V_{n\Delta}) = \frac{(n+1)!}{n_0!n_1!} \int_0^{1/2} r^{n_1}(1-r)^{n_0}\, dr = 1 - p(H^1|V_{n\Delta}). \qquad (5)$$

This can be written in terms of the jury size n and the majority Δ, using $n_0 = (n-\Delta)/2$ and $n_1 = (n+\Delta)/2$. The expectation values for the competences of the jurors, finally, are given by the following normalised integrals:

$$\begin{aligned} \mathrm{E}[q_0] &= \frac{1}{p(H^0|V_{n\Delta})} \int_0^{1/2} r^{n_1}(1-r)^{n_0+1}\, dr, \\ \mathrm{E}[q_1] &= \frac{1}{p(H^1|V_{n\Delta})} \int_{1/2}^{1} r^{n_1+1}(1-r)^{n_0}\, dr. \end{aligned}$$

If $n_1 > n_0$, then we will have that $\mathrm{E}[q_1] > \mathrm{E}[q_0]$, because on the assumption that H^0 is true a majority for H^1 is more likely if the competence q_0 is low.

Before we investigate the expression (5) in the next section, we want to address a possible criticism of the derivation of the posteriors. It may be objected that the jury vote seems to have been used twice: once for the determination of the posterior over competences, and then again for the determination of a posterior for the hypotheses based on some expected competence. But in the model above there is no such double usage. We only employ the vote to determine a probability distribution over the parameter r, which summarises the two competences q_0 and q_1. All the other probability assignments are derived from this distribution without using the data again.

4 Calculating the Posterior Probability

In the preceding section we derived a probability assignment for the hypotheses concerning Jack's guilt, under the assumption of the independence of jurors but without assuming a fixed value for the competences q_0 and q_1. The assignment is an integral expression in which both the size of the majority Δ and the jury size n play a role. In contrast, in Section 1 we presented the result by List (2004) that, on the assumption of any particular competence $q_0 = q_1 = q$, the probability of the hypotheses only depends on the majority Δ. In this section we investigate the integral of Equation (5) both analytically and numerically, thereby putting these earlier results in a new perspective.

We first give an analytic characterisation of how the probability for the hypotheses depends on jury size and absolute margin. Significantly, we retain an important consequence of the Condorcet formula, as discussed in Section 1. On the assumption that $n_1 > n_0$, or $\Delta > 0$, we have the pairwise inequality

$$p(H_r|V_{n\Delta}) \ > \ p(H_{1-r}|V_{n\Delta})$$

for all $r \in (1/2, 1]$. Hence we also have that

$$\int_{1/2}^{1} p(H_r|V_{n\Delta})\, dr \ > \ \int_{1/2}^{1} p(H_{1-r}|V_{n\Delta})\, dr\,.$$

Via the translation $1 - r = q_0$ and $r = q_1$ within $r \in (1/2, 1]$, we thus obtain the inequality $p(H^1|V_{n\Delta}) > p(H^0|V_{n\Delta})$ on condition that $\Delta > 0$.

Further, let us look at the so-called marginal likelihoods of the hypotheses H_j on the two votes $V^0_{n+1} \cap V^1_{n+2}$. These two votes effectively enlarge the jury while keeping the absolute margin Δ fixed. In Appendix A it is shown that if $\Delta = n_1 - n_0 > 0$, we have

$$p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta}) \;>\; p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta}). \tag{6}$$

Since we also have

$$\frac{p(H^1 | V_{n+2,\Delta})}{p(H^0 | V_{n+2,\Delta})} = \frac{p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta})}{p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta})} \times \frac{p(H^1 | V_{n\Delta})}{p(H^0 | V_{n\Delta})},$$

the inequality of Equation (6) entails that the odds of Jack's guilt decreases monotonically as the jury size n is increased if we hold the absolute margin $\Delta > 0$ fixed. This is in accordance with the intuitions voiced in the preceding sections, namely that the jury size affects the probability of the hypothesis.

Secondly, we investigate the limiting behaviour in order to arrive at a generalisation of Condorcet's theorem for posterior odds. The mere fact that $p(H^0 | V_{n\Delta})$ increases with the jury size for fixed $\Delta > 0$ does not yet determine the limiting value for $p(H^0 | V_{n\Delta})$ as n goes to infinity. However, as shown in Appendix B, if the absolute margin Δ is held constant in the limit, we find the asymptotic behaviour

$$\lim_{n \to \infty} p(H^0 | V_{n\Delta}) = \frac{1}{2}.$$

It is further shown that if the fractional majority, $f = \Delta / n$, is held constant in the limit, we have instead

$$\lim_{n \to \infty} p(H^0 | V_{n,nf}) = 0.$$

Finally, we show that for the latter limiting behaviour, it is not necessary for the majority to increase linearly with n. It is enough if Δ increases more quickly than $\sqrt{n}$ to ensure that $p(H^0 | V_{n\Delta})$ tends to zero.

These results are in accordance with the aforementioned intuitions on the relation between jury votes and the hypothesis voted over: if, in a large jury, we have a close call between votes in favour and votes against, this should be taken as a sign that it is hard to decide over the hypotheses, and accordingly as a reason to put less trust in how the jury has voted than may be suggested by the absolute margin.

In addition to these qualitative results, we have also done some numerical calculations with the aid of Mathematica®. The integral in Equation (5) can be written as a known transcendental function, the so-called regularised

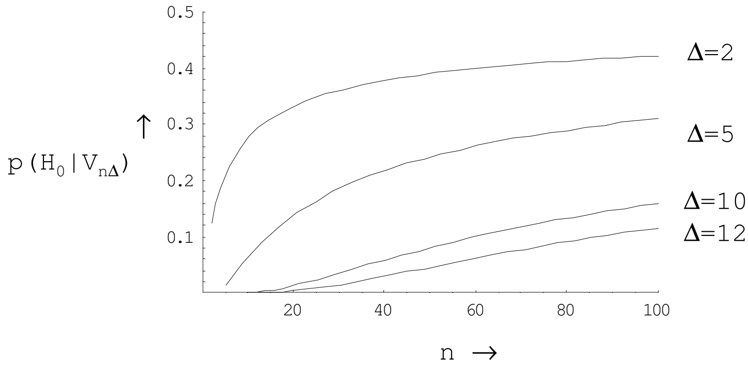

Figure 2: Graphs showing $p(H^0|V_{n\Delta})$ against jury size n for fixed values of $\Delta = 2, 5, 10,$ and 12.

incomplete Beta function. In Figure 2 we have plotted the relation between the probability $p(H^0|V_{n\Delta})$ and the jury size $n = n_1+n_0$, for various values of the size of the majority $\Delta = n_1-n_0$. These calculations illuminate the case of the two juries considered at the beginning of this paper. With a unanimous verdict of guilt in a jury of 10, $\Delta = 10$ and $n = 10$, for example, we find the probability $p(H^1|V_{10,10}) = 0.9995$. For a jury of 100 with a majority of 12 for guilt, on the other hand, we calculate a smaller probability that the jury verdict is correct, namely $p(H^1|V_{100,12}) = 0.8839$. The important point here is that the probability depends not only on the majority, as it did when we chose some fixed competence. It decreases as n is increased, and this effect may counterbalance a difference in the size of the majority.

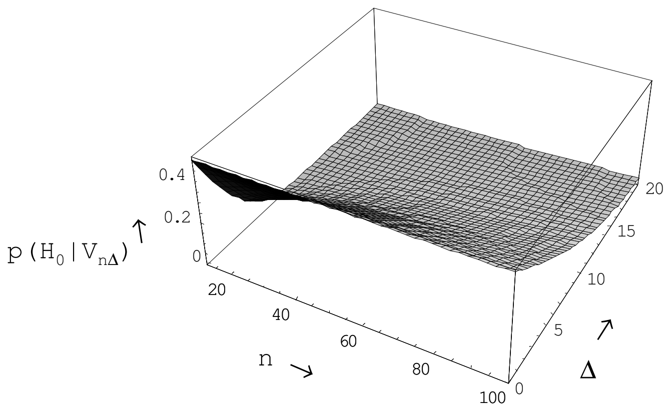

Figure 3: A graph showing $p(H^0|V_{n\Delta})$ against jury size n and majority size Δ.

In Figure 3 we see the dependence of the probability $p(H^0|V_{n\Delta})$ on the majority size Δ spelled out in more detail. Note first that for $\Delta = 0$, this probability is one-half, as it should be. Furthermore, for any fixed jury size, the probability of the hypothesis H^0 decreases with increasing majority size. And finally, for fixed majority size Δ and increasing jury size n, we can see that the probability $p(H^0|V_{n\Delta})$ slowly increases towards a half again. For very large juries, as also suggested by Figure 2, a small majority does not carry much weight.

5 Concluding remarks

Now that we have obtained these results, what can we say of the earlier result of List [2004], which stated that the probability of the hypotheses voted over only depends on the absolute margin Δ? Of course this is still a valid point under the assumption of fixed symmetric competence. But with the foregoing considerations in mind, we see that if we do not know the competences q_0 and q_1, and if we decide to learn about these competences on the basis of the jury vote, then both the absolute majority and the jury size do matter.

We conclude with some suggestions on how to develop the results of the present paper. First, we think that it is important for the practical applicability of Condorcet-style results to relax the assumption of the theorems concerning the independence of the jurors. As mentioned above, Bovens and Hartmann [2004] successfully model a jury vote with dependent jurors, and it will be interesting to see if that model can be combined with the model presented in this paper. Another way to incorporate the jury dynamics into the analysis is presented by adapting the prior, so as to make it less sensitive to almost or entirely unanimous votes. As indicated before, we might think that unanimity in small juries is due to mindless groupthink, and not a sign of a high juror competence or of truth-conducive coherence. If so, we can correct for a possible overestimation of the competence by choosing a prior over competences that is peaked around $r = 1/2$.

Perhaps a more accurate way of modelling the interaction between group members is by dropping the assumption on the likelihoods that is expressed in the left-hand equality of Equation (1) altogether. In the model presented in this paper, jury votes $V_{n\Delta}$ are characterised by the numbers n and Δ only, because the likelihoods of the hypotheses H^j only depend on these numbers. But we might also partition the space of possible jury votes differently, according to other characteristics of the votes, and employ hypotheses that have more complicated likelihood functions over that space. While this will no doubt provide interesting new insights, we can scarcely hope to attain analytic results for a model with these more involved statistical hypotheses.

An entirely different line of research concerns the possible variation of competences within the jury. Dietrich [unpublished] shows that the classical Condorcet jury theorem still holds if we suppose that the competences of jurors vary, as long as their average competence is larger than $1/2$. This raises the question whether we can also derive an expression for the posterior odds of the hypothesis on the assumption of a certain spread in the competences of the respective jurors. A suitable statistical setting for answering this question is hierarchical Bayesian modelling, in which we may suppose the juror relative competence q_{ij} to be drawn at random from a distribution of possible values for the competence. Again, analytic results may be very hard to come by, but software packages such as WinBUGS® are well equipped to investigate such models numerically.

Finally, we expect that much can be gained by applying the present insights to the discussion over the coherence measures proposed in Bovens and Hartmann [2004], and continued in Haenni and Hartmann [2006]. The reliability parameter employed there is formally similar to the competence parameter employed in the present paper. It will be interesting to see if and how we can adapt our estimations of the reliability of measurement apparatuses or witnesses from their coherence. Because the well-known impossibility result of Bovens and Hartmann relies on variability in the reliability parameter, and because in the present paper we have shown how to adapt the probability assignment over values of this parameter, we surmise that casting their impossibility result in terms of the present findings will lead to interesting new insights.

Acknowledgements

We would like to thank the XXX reading group of the Philosophy Faculty of the University of Groningen and the audience at the BSPS 2008 in St Andrews for helpful discussion.

A Relative size of the marginal likelihoods

In this appendix we prove that the marginal likelihood of the hypothesis H^0 for the combined votes $V^0_{n+1} \cap V^1_{n+2}$, given an earlier jury vote $V_{n\Delta}$ for which $\Delta > 0$, is larger than the corresponding likelihood of the hypothesis H^1. Mathematically,

$$p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta}) \ > \ p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta}). \tag{7}$$

We will do so by writing out the marginal likelihoods in terms of the likelihoods of the statistical hypotheses H_r for $r < 1/2$ and $r \geq 1/2$ respectively.

We first determine the likelihoods of the hypotheses H_r for the two votes $V^0_{n+1} \cap V^1_{n+2}$:

$$p(V^0_{n+1} \cap V^1_{n+2} | H_r \cap V_{n\Delta}) = r(1-r).$$

Recall that the hypotheses H^0 and H^1 are composed of the statistical hypotheses H_r. The likelihood of the hypothesis H^0 for the two votes $V^0_{n+1} \cap V^1_{n+2}$ is

$$\begin{aligned} &p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta}) \\ &\quad = \int_0^{1/2} p(H_r | H^0 \cap V_{n\Delta}) p(V^0_{n+1} \cap V^1_{n+2} | H_r \cap H^0 \cap V_{n\Delta})\, dr \\ &\quad = \frac{1}{p(H^0|V_{n\Delta})} \frac{(n+1)!}{n_0! n_1!} \int_0^{1/2} r^{n_1} (1-r)^{n_0}\, r(1-r)\, dr\,. \end{aligned} \tag{8}$$

in which we have used the normalisation $p(H^0|V_{n\Delta})$ because within $r \in [0, 1/2]$ we have that

$$\begin{aligned} p(H_r | H^0 \cap V_{n\Delta}) &= \frac{p(H_r \cap H^0 | V_{n\Delta})}{p(H^0|V_{n\Delta})} = \frac{p(H_r|V_{n\Delta})}{p(H^0|V_{n\Delta})} \\ &= \frac{1}{p(H^0|V_{n\Delta})} \frac{(n+1)!}{n_0! n_1!}\, r^{n_1}(1-r)^{n_0}\,. \end{aligned} \tag{9}$$

The marginal likelihood of the hypothesis H^1 for the two votes is given by a similar expression, with the difference that the integration bounds are $1/2$ and 1, and that the normalisation is $p(H^1|V_{n\Delta})$.

In order to compare the two marginal likelihoods, it will be convenient to write Equation (9) in terms of the same integration bounds, making use of the symmetry in the integral expression:

$$
\begin{aligned}
&p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta}) \\
&\quad = \frac{1}{p(H^1|V_{n\Delta})} \frac{(n+1)!}{n_0! n_1!} \int_{1/2}^{1} r^{n_1} (1-r)^{n_0}\, r(1-r)\, dr \\
&\quad = \frac{1}{p(H^1|V_{n\Delta})} \frac{(n+1)!}{n_0! n_1!} \int_{0}^{1/2} r^{n_0} (1-r)^{n_1}\, r(1-r)\, dr\,. \qquad (10)
\end{aligned}
$$

We can now compare the two marginal likelihoods by comparing the functions appearing under the integration sign. Specifically, we will investigate the expression

$$
\begin{aligned}
&p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta}) - p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta}) \\
&\quad = \int_0^{1/2} \big(p(H_r | H^0 \cap V_{n\Delta}) - p(H_r | H^1 \cap V_{n\Delta}) \big)\, r(1-r)\, dr \\
&\quad = \frac{(n+1)!}{n_0! n_1!} \int_0^{1/2} \left(\frac{r^{n_1}(1-r)^{n_0}}{p(H^0|V_{n\Delta})} - \frac{r^{n_0}(1-r)^{n_1}}{p(H^1|V_{n\Delta})} \right) r(1-r)\, dr\,, \qquad (11)
\end{aligned}
$$

the difference between the marginal likelihoods of Equations (8) and (10). If this function is positive, then the marginal likelihood of H^0 is larger than that of H^1, which is what we have set out to prove.

The expression inside the integral of Equation (11) consists of two parts. We now make some observations on the part between brackets,

$$
g(n_0, n_1, r) = \frac{r^{n_1}(1-r)^{n_0}}{p(H^0|V_{n\Delta})} - \frac{r^{n_0}(1-r)^{n_1}}{p(H^1|V_{n\Delta})}.
$$

First, because of the normalisations, $p(H_j|V_{n\Delta})$, we have

$$
\int_0^{1/2} g(n_0, n_1, r)\, dr = 0. \qquad (12)
$$

Next, if we assume that $n_1 > n_0$, Equation (6) says that $p(H^1|V_{n\Delta}) > p(H^0|V_{n\Delta})$, so that we have

$$
g(n_0, n_1, 1/2) = \left(\frac{1}{p(H^0|V_{n\Delta})} - \frac{1}{p(H^1|V_{n\Delta})} \right) \frac{1}{2^n} > 0\,. \qquad (13)
$$

Furthermore, the equation $g(n_0, n_1, r) = 0$ has two solutions in r. One is $r = 0$, the other is

$$
r^* = \frac{c}{1+c} \qquad \text{with} \qquad c = \left(\frac{p(H^0|V_{n\Delta})}{p(H^1|V_{n\Delta})} \right)^{1/\Delta}. \qquad (14)
$$

Together with Equations (11) and (13), Equation (14) entails that in the domain $r \in (0, r^*)$ we have that $g(n_0, n_1, r) < 0$ while in $r \in (r^*, 1/2]$ we have that $g(n_0, n_1, r) > 0$. Finally, with Equation (12) this entails that

$$\int_0^{r^*} |g(n_0, n_1, r)|\, dr = \int_{r^*}^{1/2} g(n_0, n_1, r)\, dr. \tag{15}$$

In other words, the entire negative contribution to the integral of Equation (12) lies in $r < r^*$, while the entire positive contribution to it lies in $r > r^*$. All this is illustrated in Figure 4.

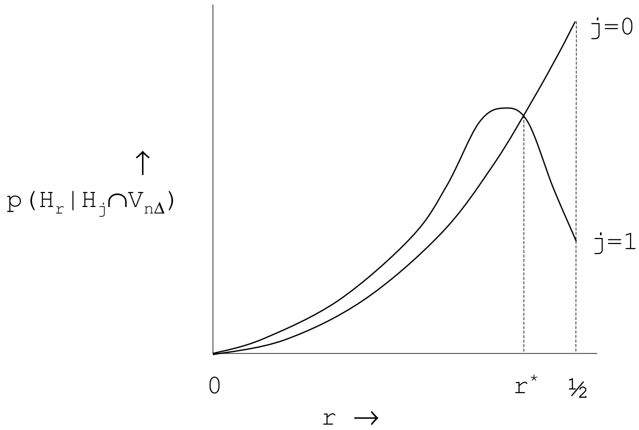

Figure 4: Graphs of the functions $p(H_r|H_j \cap V_{n\Delta})$ for $j = 0, 1$ against $r \in [0, 1/2]$. The values of n and Δ are kept fixed. As expressed in Equation (15), the two areas in between the two curves are equal.

We make one further observation on the function $r(1-r)$, namely that it is monotonically increasing in r over the domain $r \in [0, 1/2]$. Now recall that in the domain $r \in [0, r^*]$, the contribution of the integral is entirely negative. The factor with which the function $g(n_0, n_1, r)$ is multiplied over this domain is, on average, strictly less than $r^*(1-r^*)$, and the contribution to the whole integral of Equation (11) therefore has the following lower bound:

$$\int_0^{r^*} g(n_0, n_1, r)\, r(1-r)\, dr \;>\; r^*(1-r^*) \int_0^{r^*} g(n_0, n_1, r)\, dr. \tag{16}$$

In the domain $r \in [r^*, 1/2]$, on the other hand, the integral is entirely positive, and the factor with which the function $g(n_0, n_1, r)$ is multiplied is, on average, strictly more than $r^*(1-r^*)$, thus leading to a contribution with a lower bound

$$\int_{r^*}^{1/2} g(n_0, n_1, r)\, r(1-r)\, dr \;>\; r^*(1-r^*) \int_{r^*}^{1/2} g(n_0, n_1, r)\, dr. \tag{17}$$

Combining these two equations, we have a lower bound of the difference between the two marginal likelihoods covering the entire domain of r.

Hence we can determine the lower bound of the difference between the marginal likelihoods, as follows:

$$\begin{aligned}
&p(V^0_{n+1} \cap V^1_{n+2} | H^0 \cap V_{n\Delta}) - p(V^0_{n+1} \cap V^1_{n+2} | H^1 \cap V_{n\Delta}) \\
&\quad = \int_0^{1/2} g(n_0, n_1, r)\, r(1-r)\, dr \\
&\quad = \int_0^{r^*} g(n_0, n_1, r)\, r(1-r)\, dr + \int_{r^*}^{1/2} g(n_0, n_1, r)\, r(1-r)\, dr \\
&\quad > r^*(1-r^*) \left(\int_0^{r^*} g(n_0, n_1, r)\, dr + \int_{r^*}^{1/2} g(n_0, n_1, r)\, dr \right) = 0.
\end{aligned}$$

The crucial step in this derivation is of course the inequality, which is based on the two lower bounds of Equations (16) and (17). Together they establish Equation (7).

B Limiting behaviour of the probabilities

The posterior probability for H^0 conditional on the jury vote $V_{n\Delta}$ can be written as a regularized incomplete Beta function; see Abramowitz and Stegun [1964, p. 263], formulae (6.6.1) and (6.6.2). Specifically,

$$\begin{aligned}
p(H^0 | V_{n\Delta}) &= \frac{(n+1)!}{n_0! n_1!} \int_0^{\frac{1}{2}} dr\, r^{n_1} (1-r)^{n_0} \\
&= \frac{B_{\frac{1}{2}}(n_1+1, n_0+1)}{B(n_1+1, n_0+1)} \\
&\equiv I_{\frac{1}{2}}(n_1+1, n_0+1)\,. \qquad (18)
\end{aligned}$$

In this appendix we exploit certain asymptotic properties of the regularized incomplete Beta function to show that $p(H^0|V_{n\Delta})$ tends to one-half as n tends to infinity at constant Δ, but to zero if the fractional majority, $f = \Delta/n$, is held constant in the limit. It is also shown that if Δ increases more quickly than $\sqrt{n}$, $p(H^0|V_{n\Delta})$ still tends to zero in the limit of n to infinity.

Theorem 1
If $\Delta = n_1 - n_0 \geq 0$ is constant, then $I_{\frac{1}{2}}(n_1+1; n_0+1)$ tends to $1/2$ in the limit that $n = n_1 + n_0$ tends to infinity.

Proof
On changing the integration variable from r to $t = (1-2r)^2$, we find

$$I_{\frac{1}{2}}(n_1+1; n_0+1) = 2^{-n-2}\frac{(n+1)!}{n_0! n_1!}\int_0^1 \frac{dt}{\sqrt{t}}\,(1-\sqrt{t})^{n_1}(1+\sqrt{t})^{n_0}\,. \qquad (19)$$

This expression can be rewritten as

$$I_{\frac{1}{2}}(n_1+1; n_0+1) = 2^{-n-2}\frac{(n+1)!}{n_0! n_1!}\int_0^1 \frac{dt}{\sqrt{t}}\,(1-t)^{n_0}(1-\sqrt{t})^{\Delta}\,. \qquad (20)$$

The last factor in the integrand can be expanded as the finite binomial series

$$(1-\sqrt{t})^{\Delta} = \sum_{m=0}^{\Delta}\frac{\Delta!}{m!\,(\Delta-m)!}(-1)^m\,t^{\frac{m}{2}}\,,$$

and this allows the evaluation of the integral, term for term:

$$I_{\frac{1}{2}}(n_1{+}1; n_0{+}1) = 2^{-n-2}\frac{(n+1)!}{n_0!\,n_1!}\sum_{m=0}^{\Delta}\frac{\Delta!}{m!\,(\Delta-m)!}(-1)^m\frac{\Gamma(\frac{m+1}{2})\Gamma(n_0+1)}{\Gamma(n_0+\frac{m+3}{2})}\,. \qquad (21)$$

The Stirling expansion, namely

$$\Gamma(n) = (n-1)! = \sqrt{2\pi}n^{n-\frac{1}{2}}\,\mathrm{e}^{-n}[1+O(\tfrac{1}{n})]$$

is now used to give the asymptotic expressions

$$\begin{aligned}\frac{(n+1)!}{n_0!\,n_1!} &\sim 2^{n+1}\sqrt{\frac{n}{2\pi}}\\ \frac{\Gamma(n_0+1)}{\Gamma(n_0+\frac{m+3}{2})} &\sim n_0^{-\frac{m+1}{2}} = \left(\frac{2}{n-\Delta}\right)^{\frac{m+1}{2}}\,.\end{aligned}$$

On inserting these forms into Equation (21), we find that $I_{\frac{1}{2}}(n_1+1; n_0+1)$ is asymptotically equivalent to

$$\frac{1}{2\sqrt{\pi}}\sqrt{\frac{n}{n-\Delta}}\left[\Gamma(1/2) - \Delta\Gamma(1)\sqrt{\frac{2}{n-\Delta}} + \ldots + (-1)^{\Delta}\Gamma(\tfrac{\Delta+1}{2})\left(\frac{2}{n-\Delta}\right)^{\frac{\Delta}{2}}\right]$$

All the terms in the square braces, except for the first one, vanish in the limit of large n, and there is only a finite number of these terms. So only the first term survives, and since $\Gamma(1/2) = \sqrt{\pi}$, we have thereby proved indeed that $I_{\frac{1}{2}}(n_1+1; n_0+1)$ tends to $1/2$ in the limit as n tends to infinity.

Theorem 2
If $\Delta \sim n^\beta$ for large n, with $1/2 < \beta \leq 1$, then $I_{\frac{1}{2}}(n_1+1; n_0+1)$ tends to 0 in the limit.

Proof
Recall that we can rewrite Equation (18) as Equation (19). The latter can also be rewritten as

$$I_{\frac{1}{2}}(n_1+1;n_0+1) = 2^{-n-2}\frac{(n+1)!}{n_0!n_1!}\int_0^1 \frac{dt}{\sqrt{t}}\,(1-t)^{n_1}(1+\sqrt{t})^{-\Delta}\,, \qquad (22)$$

and we now split the integral into two pieces, corresponding to $0 < t < \varepsilon^2$ and $\varepsilon^2 < t < 1$, where ε will be specified in a moment. Clearly,

$$\int_0^{\varepsilon^2} \frac{dt}{\sqrt{t}}\,(1-t)^{n_1}(1+\sqrt{t})^{-\Delta} < \int_0^{\varepsilon^2}\frac{dt}{\sqrt{t}} = 2\varepsilon$$

whereas

$$\begin{aligned}\int_{\varepsilon^2}^1 \frac{dt}{\sqrt{t}}\,(1-t)^{n_1}(1+\sqrt{t})^{-\Delta} &< (1+\varepsilon)^{-\Delta}\int_0^1\frac{dt}{\sqrt{t}}\,(1-t)^{n_1} \\ &= (1+\varepsilon)^{-\Delta}\frac{\Gamma(\frac{1}{2})\Gamma(n_1+1)}{\Gamma(n_1+\frac{3}{2})}\,.\end{aligned}$$

We insert the last two inequalities into Equation (22) and also employ the Stirling expansion, as in the proof of Theorem 1. This yields

$$I_{\frac{1}{2}}(n_1+1;n_0+1) < \sqrt{\frac{n}{2\pi}}\varepsilon + \sqrt{\frac{n}{n+\Delta}}(1+\varepsilon)^{-\Delta}\,.$$

Now choose $\varepsilon = n^{-\alpha}$ and put $\Delta \sim n^\beta$, thereby obtaining

$$I_{\frac{1}{2}}(n_1+1;n_0+1) < \frac{1}{\sqrt{2\pi}}\,n^{\frac{1}{2}-\alpha} + (1+n^{-\alpha})^{-n^\beta}\,.$$

Now $(1+n^{-\alpha})^{n^\alpha}$ tends to e in the limit of large n, so we obtain

$$I_{\frac{1}{2}}(n_1+1;n_0+1) < \frac{1}{\sqrt{2\pi}}\,n^{\frac{1}{2}-\alpha} + \exp[-n^{\beta-\alpha}]$$

asymptotically. For any $\alpha > 1/2$, the first term above vanishes asymptotically, and for any $\beta > \alpha$, so does the second term. Hence for any $1/2 < \beta \leq 1$ we have shown that $I_{\frac{1}{2}}(n_1+1;n_0+1)$ tends to 0 in the limit as n tends to infinity.

Corollary

If $f = \frac{\Delta}{n}$ is constant, then $I_{\frac{1}{2}}(n_1 + 1; n_0 + 1)$ tends to 0 in the limit that $n = n_1 + n_0$ tends to infinity.

Proof

This follows immediately by taking the special case $\beta = 1$ in Theorem 2.

References

Abramowitz, M. and I. A. Stegun [1964] : *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, New York: Dover.

Bovens, L. and Hartmann, S. [2003] : *Bayesian Epistemology*, Oxford: Oxford University Press.

Dietrich, F. [2008] : 'The premises of Condorcet's jury theorem are not simultaneously justified', unpublished manuscript.

Haenni, R. and S. Hartmann [2006] : 'Modeling partially reliable information sources: a general approach based on DempsterShafer theory', *Information Fusion*, **7**, pp. 361–79.

List, C. [2004] : "On the Significance of the Absolute Margin", *British Journal for the Philosophy of Science*, **55**, pp. 521–44.